MOLECULAR PHYLOGENY OF THE GENUS *VICIA* L. (FABACEAE) BASED ON ITS2 AND COX1 HOUSEKEEPING GENES

Eyerusalem Sija Shehbala

Master thesis

Erasmus Mundus Master Program in Plant Breeding
(emPLANT)

University of Helsinki, Department of Agricultural Sciences

May 2021

UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of Agriculture and Forestry	Department of Agricultural Sciences

Author
Eyerusalem Sija Shehbala

Title
MOLECULAR PHYLOGENY OF THE GENUS *VICIA* L. (FABACEAE) BASED ON ITS2 AND COX1 HOUSEKEEPING GENES

Subject
Plant breeding

Level	Month and year	Number of pages
M.Sc. Thesis	May 2021	30

Abstract

Genus *Vicia* comprises more than 190 species that chiefly grow in the temperate area of Europe and Asia as well as some parts of tropical Africa, North America, and South America. Based on inflorescence length and the presence of nectariferous spots on the stipules, the genus is divided into two subgenera *Cracca* (*Vicilla*) and *Vicia*, consist of 17 and 9 sections, respectively. The most widely known species is faba bean (*V. faba*), also referred to as poor man's meat. However, the close relatedness of *V. faba* from section *Faba* with section *Narbonensis* is still uncertain while *V. hirsuta* placed in section *Cracca* although morphological and biochemical evidence support its separation from *Cracca*. Moreover, there is also the issue of further dividing the genus by separating *Ervum* as the third subgenus. Therefore, this work aimed to examine the evolutional relationship among 24 selected species of genus *Vicia* by using housekeeping genes, nuclear ribosomal internal transcribed spacer (ITS2) and mitochondrial cytochrome c oxidase subunit I (Cox1). The plant DNA was extracted and amplified with gene-specific primers followed by PCR product clean up by using GeneJET. The cleaned DNA fragments were cloned with a pGEM-T vector and sequenced. The sequence alignment was performed using CLUSTAL W and MAGA-X to construct maximum likelihood phylogenetic trees. The result showed that most of the species were monophyletic and there were few paraphyletic groupings. Species in section *Narbonensis* were in the last part of the tree whereas *V. faba* was placed in the upper part far from those species. *V. hirsuta* formed a clade with section *Narbonensis* rather than with section *Cracca*. Similarly, section *Ervum* grouped with species of section *Cassubicae* and *Panduralae* instead of being separated from the rest of the species. In conclusion, this work showed the molecular phylogenetic tree of 24 selected species of genus *Vicia* which are able to provide more acceptable details to classify as compared to previous *Vicia* morphological phylogenetic tree. ITS2 was a more informative tool than Cox1 for dealing with phylogenetic relationships.

Keywords
Vicia, Molecular Phylogenetic, ITS2, Cox1, Maximum likelihood

Where deposited
Master's Program in Agricultural Sciences, Department of Agricultural Sciences

Supervisor
Adjunct Prof. Fred Stoddard

Table of Contents

1. Introduction

Legumes (Fabaceae) are the third biggest flowering (angiosperm) family and the second most nutritional important to human after the cereals (Poaceae). Legumes are a good source of protein and affordable, so they are considered as suitable replacements of animal products as source of protein in the human diet. Their secondary metabolites can be used as a raw material for different industries such as pharmaceuticals to produce medicine and fortified food, chemical industry to produce environmentally friendly pesticides as well as in the production of biofuels. In addition, their symbiotic relationship with nitrogen-fixing bacteria in the roots increases their value in agricultural production and productivity. They also a play major role as a source of animal feed (Raveendar et al., 2015).

Vicia L. is a genus belonging to the family Fabaceae that has more than 190 species that mainly grow in the temperate areas of Asia, Europe and tropical Africa as well as North America and South America (van de Wouw et al., 2001; Abozeid et al., 2018). Currently, it is believed that Europe and Asia account for approximately 110 species while around 15 species are endemic to Africa. North America and South America each hosts about 20 species (Roze & Rūrāne, 2013). Faba bean *(Vicia faba* L*)* is one of the most important and well-known species of the genus. It is mainly used as human food especially in West Asia and North Africa and also used as animal feed elsewhere in the world (van de Wouw et al., 2001). Moreover, in Europe the use of faba bean as food has been increased. Specially in Nordic country several companies are promoting health and environmental benefit of plant based protein (Gold and Green 2020, Verso Food 2021).

Genus *Vicia* can contribute to filling current world needs regarding agricultural crop production and animal feed (van de Wouw et al., 2001). However, many members of this genus contain toxic and antinutritional compounds that are harmful for human and animal health (Berger et al., 2003; Lahuta et al., 2018). Studying and understanding taxonomic relationships between the important crops and their close relatives is essential to bring the desired characteristics from wild relative to our target plant through different breeding methods (Abozeid et al., 2018; Caputo et al., 2013). The background of taxonomic classification of genus *Vicia* is relatively stable and based on Kupicha's (1976) classification with minor modification (Hanelt & Mettin, 1989).

Genus *Vicia* is characterized by inflorescence length and the presence or absence of nectaries in the stipules. The two features are used to classify the genus into two subgenera. The first

and largest is *Cracca* (*Vicilla)* which contains 17 sections including major forage species. The second one is *Vicia* which has 9 sections covering important agricultural legumes (van de Wouw et al., 2001). Even though the current classification is widely accepted, still there is disagreement among taxonomists regarding the relation of *V. faba* with its close relatives as well as the placement of different species in a particular section (Caputo et al., 2013).

Morphology based classification of *Vicia* species has been the reason that still there is unresolved and floating taxonomic group among the species. This happens because most of the species have intermediary structural characters that make it almost impossible to distinguish them without any molecular aid (Raveendar et al., 2015). Phylogenetic categorization based on the genomic DNA sequence provides more knowledge and widens the scope of understanding by giving the whole perspective regarding the study target. Furthermore, the extracted information from a DNA sequence can be used to study genetic evolution between and within species (Frediani et al., 2004).

2. Literature review

2.1 Morphology based classifications of *Vicia*

The role of phylogenetic information in different field of study has become more common and useful in the last 20 years. Having diversified phylogenetic data is becoming more important in biology conservation and other biodiversity questions (Janssens et al., 2020). Phylogenies can solve problems raised from relatedness in biological communities. However, understanding the whole evolution of an organism requires combinations of several data types including phenotypes, genomic data and geographic distribution. Specially in the plant community, there is not enough evidence related to species to produce phylogenies, so that biologists and plant ecologists are forced to use common methods to generate trees that will work for their intended aim (Qian & Jin, 2016); Bouckaert et al., 2019).

Vicia is taxonomically divided into two subgenera, *Vicia* and *Cracca* (*Vicilla*) based on the morphological categorization of Kupicha (1976) according to the relative length of the inflorescence and presence or absence of nectariferous spots on the stipules (Leht & Jaaska, 2002; Leht, 2005; Leht, 2009). The subgenus *Vicia* is a small and aggregated taxon. The species in the subgenus share a number of unique features and it can be easily distinguished. Due to those reasons, there have been more studies sas compared to *Cracca* (*Vicilla*) (Leht, 2005; Leht, 2009).

The subgenus *Vicia* is characterized by the presence of nectariferous spots on the stipules as well as with short inflorescence. There is considerable morphological difference among the members. The number of sections in this subgenus is stated differently by different authors from 5 to 9. According to different characterizations used to group the species, there are 20 to 80 species in the subgenus *Vicia*. Most species are self-pollinated annuals. There are also cross-pollinated or partially cross-pollinated species such as *V. grandiflora* and *V. faba*. Species of sections *Sepium* (Radzhia) and *Lathyroides* (Buchenau) are cross-pollinated perennials. Several agricultural important crops including faba bean (*V. faba),* narbon vetch (*V. narbonensis* subsp. *narbonensis*), and field vetch (*V. sativa* subsp. sativa) belong to this subgenus (Hanelt & Mettin, 1989; Leht & Jaaska, 2002)

Cracca is the correct name of the second subgenus rather than *Vicilla* (Leht, 2009). It is the larger subgenus and contains about 140 to 160 species. Due to its lack of uniformity and common morphological and other biochemical characterizations in a major part of this subgenus, it is difficult to give stable and precise categorization. The large number of *Cracca* members are characterized by the following phenotypic traits: perennial, many flowered, long-peduncled, simple stipules, and more primary scale leaves on main stem (Hanelt & Mettin, 1989). *Cracca,* unlike subgenus *Vicia*, has not received much modification to its 17 sections. The changes done by recent work of taxonomist from Kupicha (1976) are basically in the number of subgenera, sections and the categorization of some species in section for example, *V. hirsuta* (L.) S. F. Grey, *V. sylvatica* L. and *V. biennis* L placement is different from Kupicha (1976). Most forest *Vicia* species of Eurasia belongs to subgenus *Cracca* (Hanelt & Mettin, 1989; Leht, 2005).

Leht's (2009) study placed the *Ervoid* group in a separate subgenus *Ervum* (Figure 1) which is argued differently from Kupicha (1976). The supporting idea in is that the *Ervoid* group forms a monophyletic clade outside both subgenus *Vicia* (I) and *Cracca* (II). Subgenus *Ervum* has 4 sections, namely *Ervum*, *Ervilia*, *Ervoides* and *Lentopsis*.

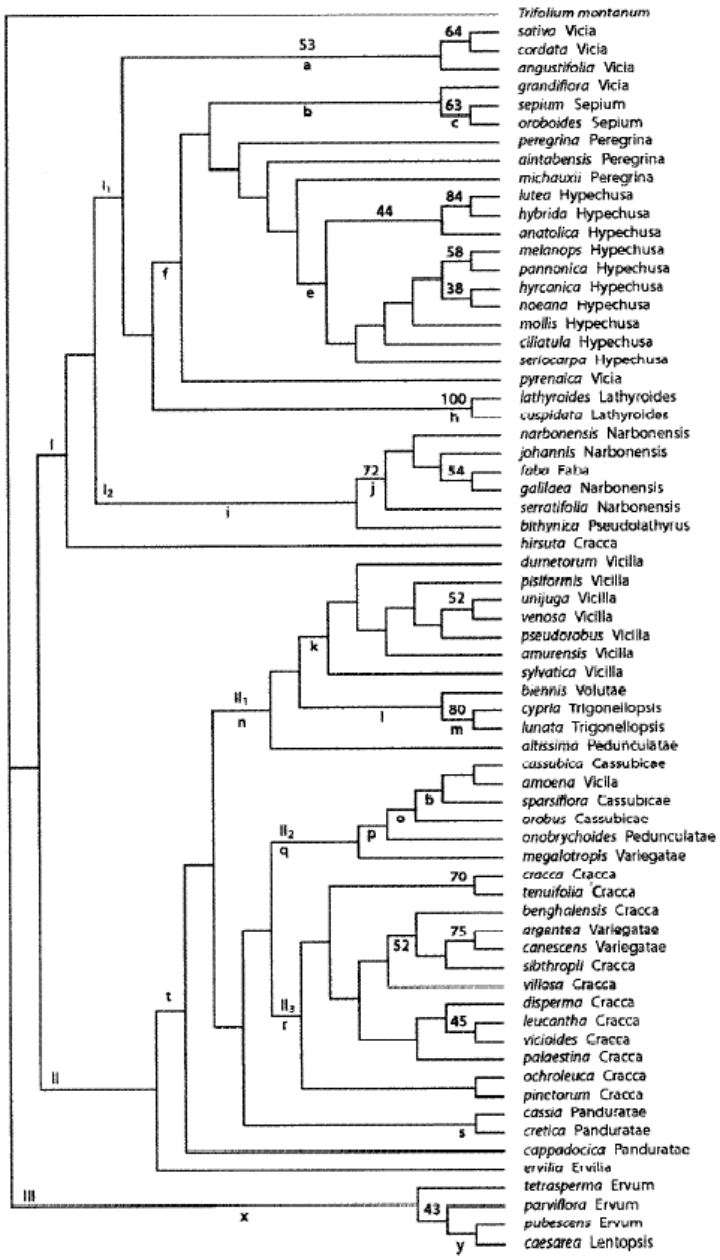

Figure 1. Phylogeny of genus *Vicia* based on morphology (Adopted from Leht, (2009)).

2.2 Biochemical and Molecular based classifications of *Vicia*

Isoenzyme analysis is commonly used in taxonomy of groups with very broad as well as diversified morphology. More often the result can be implemented to combine different species or to apply segregation among the members in certain genus. It is simply done based on the number of a band in the electrophoresis gel as well as allelic frequency and ratio can be used to draw a hierarchical hypothesis (Micales & Bonde, 2017). From subgenus *Cracca* (*Vicilla*) the cladistic and phenetic relationships of 27 species were studied by using isoenzyme composition variation pattern at 15 loci. The analysis found that there were both polyphyletic

and monophyletic categories in different sections of this subgenus. For example, sections *Cracca*, *Ervum*, *Pedunculatae* and *Lenticula* formed monophyletic groups, whereas section *Oroboideae* and *Panduratae* were polyphyletic. On the other hand, the isoenzyme result did not separate section *Ervum* into a third subgenus. Hence, it formed paraphyly with other members of subgenus *Cracca* (Jaaska, 2005).

Jaaska (1997) used polyacrylamide gel electrophoresis (PAGE) to separate isoenzymes of ten enzymes and allozymes that were genetically different and coded by 17 loci of 21 species of subgenus *Vicia*. According to the cladistic analysis of allozyme data, there were two monophyletic classes with two monophyletic subclades for each. The first subclade of the first class contained *Vicia*, *Sepium*, *Pseudolathyrus* and *Lathyroides* related species. In the second subclade of the first class were species from section *Peregrinae*. The first subclade of the second monophyletic class comprised section *Hyperchusa* species such as *V. hybrida, V. pannonica, V. anatolica* and *V. melanops,* along with section *Narbonensis*. The morphology of *Narbonensis* resembles that of faba bean. The second subclade of the second monophyletic class was paraphyletic and consisted of *V. hyrcanica* and *V. lutea* from section *Hyperchusa* (Jaaska, 1997; Potokina et al., 1999; Leht & Jaaska, 2002). The isoenzyme analysis evidence showed that section *Narbonensis* is not the closest relative to the cultivated faba bean as is often stated from their morphological resemblance (Jaaska, 1997).

Restriction site variation and repeating length of *Vicia* ribosomal DNA (rDNA) were assessed in 49 species of the genus. DNA restriction endonuclease fragment analysis showed that there was a significant difference between *V. faba* and members of *Narbonensis* species complex (Raina & Ogihara, 1995; Jaaska,1997). Moreover, the genetic relation between the *Narbonensis* species complex and *V. bithynica* with *V. faba* was distant and they do not share synapomorphic substitutions; rather, *V. faba* was more related to section *Peregrinae* (Shiran et al., 2014).

2.3 Phylogeny using nuclear gene ITS

The internal transcribed spacer (ITS) is located between 16S-5.8S-26S regions and has three different regions (Figure 2), the conserved parts being ITS1 and ITS2 and there is the exon part between of 5.8S. The total length of ITS is different depending on the species of plant, in angiosperms it could be from 500 to 750 bp whereas in other seed plants it is much longer, between 1500 and 3500 bp (Poczai & Hyvönen, 2010). Nuclear internal transcribed spacer (ITS) regions sequence has been commonly used as a DNA marker in plant phylogenetics

(Kress et al., 2005; Cowan et al., 2006; Cheng et al., 2016). ITS is highly variable even among closely related species, so it is very suitable to study taxonomies. It has a very conserved region that can be used to assess ancestral diversification and roots (Tippery & Les, 2008; Miranda et al., 2010).

ITS1 and ITS2 has been very useful tools for the taxonomic classification of different legumes (Wojciechowski et al., 2004; Shiran et al., 2014). In addition to phylogeny, the structural data of family tree from ITS can give useful information in order to draw a hypothesis to understand ancestral root (Tippery & Les, 2008).

Internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA were analyzed for comparison of stylar features of 48 taxa of Fabaceae. According to stylar shapes and hairy patterns, it was possible to distinguish the monophyletic and paraphyletic relation among the taxa. There were four types of style: Dorsiventrally compressed and evenly hairy all round (De-type), terete and evenly hairy all round (Te-type), laterally compressed and evenly hairy all round (Le-type), dorsiventrally compressed and abaxially tufted (Dabt-type). Species from East Asia including members of section *Amurense* and European species such as *V. pisiformis* formed a clade with subgenus *Cracca* species from East Asia since they have De-type style. *V. hirsuta* and *V. vicioides* (Desf.) were characterized by Te-type styles. All species-related to section *Cracca* have Le-type styles (Choi et al., 2006).

Phylogenetic relations of 49 species of genus *Vicia* were presented by examining ITS1-5.8S-ITS2 region of the 18S-28S nrDNA sequences. The result showed that there was considerable distance between the Narbonensis species complex (NSC) and V. *faba* as well as with *V. bithynica*. Instead, both *V. faba* and *V. bithynica* formed close groupings with sections *Peregrinae* and *Vicia,* respectively. In addition, except for the placement of *V. lathyroides* in section *Vicia* and of both V. *vicioides* and *V. hirsuta* in section *Cracca*, the current section placement was accepted. ITS analysis showed that subsp. *johannis* from NSC possess six synapomorphic substitutions that gave support for it to be ranked as a species. This result also found that section *Hypechusa* was polyphyletic and paraphyletic (Shiran et al., 2014).

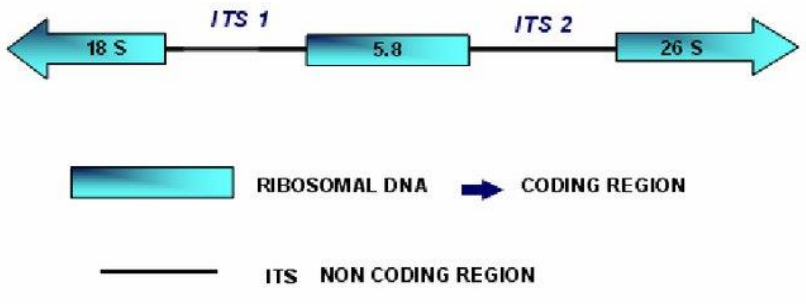

Figure 2. Genomic organization of the18S rDNA, ITS1, 5.8S rDNA, ITS2 and 26S rDNA (adapted from (Alcoba-Flórez et al., 2007))

2.4 Phylogeny using Mitochondrial gene Cox1

Plant phylogenetic analysis will be more useful if the analysis includes the data extracted from mitochondrial, plastid, and nuclear genomes together and separately by using available techniques and methods (Barkman et al., 2000). Unlike chloroplast genomes, plant mitochondrial DNA (mtDNA) experiences frequent changes including rapid genomic modification, active cooperation with nucleus by transferring genes as well as integration with unknown DNA in the nucleus and changes of gene continuity in introns or exons. On the other hand, at the sequence level mtDNA is the most slowly evolving genome of the three plant genomes (Knoop, 2004).

Mitochondrial cytochrome oxidase subunit I (cox1) sequence has been accepted and used for animal bar coding. However, because of its slow gene evolution in plants, it is not considered advisable to use it. Those slow changes at sequence level will not be useful to distinguish closely related species or generally at the species level (Kress et al., 2005; Fazekas et al., 2008). Despite this, Knoop (2004) stated that the slow evolution of gene and sequence still contains important phylogenetic information that can be used to look back at the ancestry and understand phylogeny. Phylogeny based on cox1 and atpA provided strong evidence in the process of clarifying the relationship between gnetales to angiosperms. The analysis showed Gnetales and conifers are close relative by rejecting the hypothesis of gnetales are sister to angiosperms (Bowe et al., 2000).

3. Research Objective

This study aimed at providing useful information regarding the phylogenetic relationship of genus *Vicia* at the species level from a different location. For this purpose, 24 selected species of the genus were assessed using a nuclear ribosomal internal transcribed spacer (ITS2) and mitochondrial cytochrome c oxidase subunit I (Cox1) as a study tool. The information derived

from the molecular phylogenetic tree will significantly help to resolve the disagreement on some of the species which are not still categorized in the right section. Moreover, molecular phylogenetic classification can be used with morphological information in order to understand the genus better.

4. Material and Methods

4.1 Plant material

Based on their geographical distribution, 24 species of the genus *Vicia* were selected for this work. Aurora, ILB938 and Melodiei are three cultivar of *V. faba* were used as a reference for ITS2 since faba bean are relatively well studied compared to other *Vicia* species (Table 1). The leaf material was taken from the work collection of Ph.D. candidate Laura Vottonen's at the University of Helsinki, Finland.

Table 1. Plant species used in this work with their code in GenBank and the country where planting material was collected as well as with their current section they are placed.

No_	Section	Name of species	Country	Code
1	Vicilla	Amurensis	China	VIC 82
2	Pseudolathyrus	Bithynica	Greece	VIC 12
3	Panduratae	Cassia	Turkey	VIC 5245
4	Cassubicae	Cassubica	Poland	VIC 42
5	Cracca	Cracca	Spain	VIC 72
6	Ervilia	Ervilia	Syria	ERV 111
7	Faba	Faba		
8	Narbonensis	Galilaea	Israel	NAR 44
9	Vicia	Grandiflora	Hungary	VIC 736
10	Cracca	Hirsuta	India	VIC 861
11	Vicia	Lathyroides	Slovakia	VIC 731
12	Hypechusa	Melanops	France	VIC 475
13	Peregrinae	Michauxii	Armenia	VIC 831
14	Narbonensis	Narbonensis	Italy	NAR 35
15	Cracca	Ochroleuca	Yugoslavia	VIC 73
16	Vicilla	Orobus	Norway	VIC 67
17	Peregrinae	Peregrina	Kazakhstan	VIC 5258
18	Ervum	Pubescens	Portugal	VIC 478
19	Vicia	Sativa	Tunisia	VIC 1200
20	Sepium	Sepium	Finland	VIC 56
21	Narbonensis	Serratifolia	Hungary	NAR 142
22	Ervum	Tetrasperma	UK	VIC 726
23	Vicilla	Unijuga	Russia	VIC 5146
24	Cracca	Villosa	Hungary	VIC 723

4.2 DNA extraction

Plant DNA extraction was done according to the cetyltrimethylammonium bromide (CTAB) protocol of Doyle & Doyle (1990) with a slight modification in the CTAB preparation by adding 2.8 µl proteinase K (Thermo Fisher Scientific Baltics UAB, Vilnius, Lithuania) per sample. The leaf sample was collected and frozen at -20°C. For DNA extraction those were frozen again with liquid nitrogen and ground in an MM300 TissueLyser Lab Vibration Mill Mixer (Retsch, Haan, Germany). A total of 0.9 ml of CTAB lysis buffer with mercaptoethanol was added to the ground leaf and vortexed to mix it well. The sample was then incubated in the 60°C heat block for 30 min. After 30 min, it was run in the centrifuge for 3 min at 8000 x g then 750 µl of supernatant was transferred into 2 ml tube with 750 µl chloroform-isoamyl alcohol and again centrifuged for 3 min. Then carefully, 650 µl of the upper part of the sample was taken into a fresh 2 ml tube to mix it with 650 µl chloroform-isoamyl alcohol, vortexed, and centrifuged for 3 min. To precipitate, 450 µl of the sample was transferred to a new 1.5 ml tube and then mixed with 300 µl ice-cold isopropanol. Later, it was incubated at room temperature for 30 min, followed by centrifugation for 10 min in 4 °C at full speed. When centrifugation finished, the pellet was washed with 500 µl of 70% ethanol. Subsequently, the samples were centrifuged for 5 min in the 4 °C centrifuges at full speed and later the ethanol removed. The sample was put at room temperature for 10-15 min to let the residual ethanol evaporate.

Next, the samples were resuspended with 50 µl RNAse plus TE- buffer and incubated in the heat block for 30 min at 37 °C. A total of 400 µl TE-buffer and 500µl chloroform-isoamyl alcohol was added to the sample to be centrifuged for 5 min at 10 000 rpm. A volume of 450 µl from the upper phase of the sample was transferred to a new 1.5 ml tube. The sample was then kept in the room for a few min to remove chloroform by evaporation. Finally, 45 µl of 3 M NaCl and 1 ml ice-cold 99% ethanol was added, and the samples were kept in -20 °C freezer overnight. On the next day, the sample was centrifuged at full speed for 15-20 min at + 4°C. Following centrifugation, the liquid was removed from the pellet and washed with 100 µl of 70 % ethanol. Lastly, it was run for 5 min at full speed in a + 4°C centrifuge and air-dried. Afterward, the sample eluted with 25 µl TE-buffer then stored in a -20 °C freezer. The quality and quantity of the extracted DNA was evaluated with 1 µl in the NanoDrop™ Lite Spectrophotometer (ThermoFisher Scientific Inc, Wilmington, DE, USA).

4.3 Amplification of DNA

ITS2 and Cox1 amplification reaction was done in a total volume of 50 µl polymerase chain reaction (PCR) mix that contained 2 µl DNA, 9 µl 10 x Taq buffer, 1 µl of 10 mM dNTPs,1 µl Taq DNA polymerase, and 34.5 µl MilliQ water with 1.25 µl of their perspective forward and reverse primer (Table 2). The PCR reaction for ITS2 was started with a 4 min heating at 94°C, followed by 34 cycles consisting of 94°C denaturation for 30 s, primer annealing at 55 °C for 40 s, and extension at 72°C for 1min. Reactions ended with a final step of 10 minutes at 72 °C. In order to amplify a Cox1 region the following program adjusted in the PCR machine, 94 °C for 2 minutes, 94 °C for 30 s with 35 cycles, 51 °C for 40 s, 72 °C for 1 minute with a final extension at 72 °C for 10 min. The 5 µl PCR product was then subjected to 2.5 % agarose gel with ethidium bromide (Figures 3 and 4). The rest of the PCR products was purified according to the manufacture's instruction of the GeneJET PCR Purification Kit (Thermo Fisher Scientific, California, USA).

Table 2. List of primer used for amplification of ITS2 and Cox 1.

Gene name	Primer name	Sequence 5′ to 3′	Z
ITS2	ITSp3F	YGACTCTCGGCAACGGATA	(Cheng et al., 2016)
	ITSu4R	RGTTTCTTTTCCTCCGCTTA	
Cox 1	Cox1-42F	GGATCTTCTCCACTAACCACAAA	(Fazekas et al., 2008)
	Cox1-ajf699R	′CCGAAAGAGATGCTGGTATA	

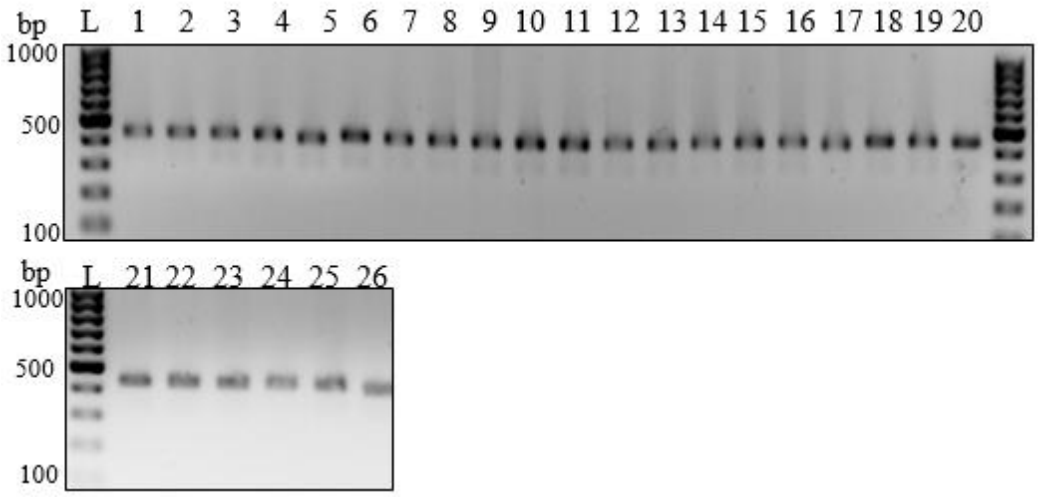

Figure 3. Electrophoresis gel shows that the amplification of the ITS2 region of the selected species from *Vicia* by using the primer listed in table 2. The size and the single band were as it was expected. 1.Amurensis, 2. Faba (aurora), 3. Bithynica, 4. Cassia, 5. Cassubica, 6. Cracca, 7. Ervilia, 8. Galilaea, 9. Grandiflora, 10. Hirsuta, 11. Faba (ILB938), 12. Lathyroides, 13. Melanops, 14. Faba (Melodie), 15. Michauxii, 16. Narbonensis, 17. Ochroleuca, 18. Orobus,

19. Peregrina, 20. Pubescens, 21. Sativa, 22. Sepium, 23. Serratifolia, 24. Tetrasperma, 25. Unijuga, 26. Villosa.

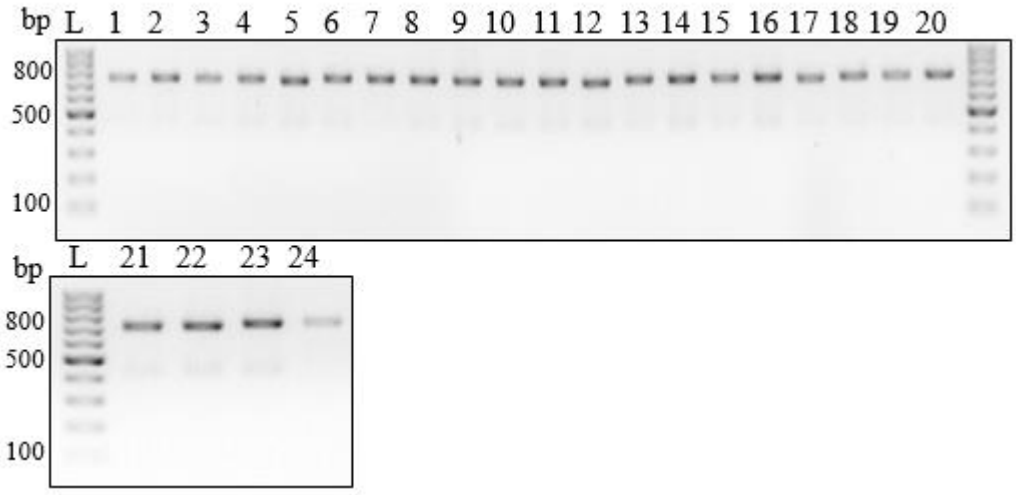

Figure 4. Electrophoresis gel shows that the amplification of the Cox1 region of 24 selected species with their specific primer (Table 2). The expected size and the single band can be seen. 1.Amurensis, 2. Bithynica, 3. Cassia, 4. Cassubica, 5. Cracca, 6. Ervilia, 7. Galilaea, 8. Grandiflora, 9. Hirsuta, 10. Lathyroides, 11. Melanops, 12. Faba (Melodie), 13. Michauxii, 14. Narbonensis, 15. Ochroleuca, 16. Orobus, 17. Peregrina, 18. Pubescens, 19. Sativa, 20. Sepium, 21. Serratifolia, 22. Tetrasperma, 23. Unijuga, 24. Villosa.

4.4 Vector cloning and Sequencing

Purified PCR products were ligated with pGEM-T vector and cloned according to the protocol in pGEM®-T vector system I (Promega, Madison, USA). To get recombinant plasmid, the product of cloning grew into Luria–Bertani plates that contained ampicillin as a selection marker. Only the positive colony was able to grow on a plate with ampicillin. Consequently, colony PCR was performed and checked in 1 % Agarose gel containing ethidium bromide. Based on the PCR confirmation, DNA was extracted by following the manual instruction in the GeneJET Plasmid Miniprep kit (Thermo Fisher Scientific, California, USA). DNA sequencing operated in Eurofins Genomics Germany GmbH (Eurofins, Ebersberg, Germany) and both ITS2 and Cox1 primers were used.

4.5 Data analysis

The result of the DNA sequence from both ITS2 and Cox1 was checked in VecScreen of National Center for Biotechnology Information (NCBI) for any vector contamination. After substantiating the contamination, trimming was performed on CLC Genomics Workbench (Version 21.0.3, QIAGEN, Aarhus, Denmark). Individual DNA alignment was implemented using CLUSTAL W (Thompson et al., 1994). Phylogenetic trees from both genes were constructed from the concatenation of ITS2 and Cox1 through PhyloSuite (Zhang et al., 2020).

The substitution model selection and construction of a phylogenetic tree were managed via MEGA_X (Kumar et al., 2018). Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best (Nei & Kumar, 2000). During the selection for the best-fitting model, the default setting is used in MEGA-X. Combined phylogenetic trees of ITS2 and Cox1 were analyzed under maximum likelihood. A thousand replications of bootstrap were performed for this purpose (Felsenstein, 1985).

5. Results

The primary objective of this study is to set the first step in the process of understanding phylogenetic relationship of genus *Vicia* at species level from different place. In order to accomplish this purpose, molecular phylogeny based on nuclear and mitochondrial DNA regions amplification and sequencing was performed, then sequence alignment was completed in both ITS2 and Cox1 DNA sequences. The result was used to construct neighbor-joining, maximum likelihood, and maximum parsimony. However, the phylogenetic tree from neighbor-joining, and maximum parsimony had low bootstrap support because the aligned DNA sequence had poor similarity. Therefore, in this work only the phylogenetic tree from maximum likelihood analysis is included since it is the best fit method according to the nucleotide substitution model selection.

5.1 ITS2 Phylogenetic Analyses

The lowest BIC scores (Bayesian Information Criterion) were 3453.471. Maximum likelihood (ML) analyses on 514 bp of ITS2 sample were based on the following substitution model, K2 (Kimura 2-paramete) for which maximum likelihood value (—ln L) = 1498.906 and all nucleotide frequencies were 0.250. Rates of base substitutions of A—T, A—C, T—G, C—G were 0.038 whereas A—G and T—C = 0.174. This analysis involved 26 nucleotide sequences. All positions containing gaps and missing data were eliminated. There was a total of 345 positions in the final dataset and evolutionary analyses were conducted.

There were two groups that shared one common ancestor. *V. cassubica*, *V. orobus* and *V. cassia* from section *Cassubica* and *Panduralae* from subgenus *Cracca* formed one clade that was supported by 70% and 40% bootstrap, respectively. At the same time, this group formed a monophyletic clade with species of section *Ervum* including *V. pubescens* and *V. tetrasperma* and was supported by 22 % bootstrap, while the monophyletic clades of section *Ervum* had 94 % bootstrap support. *V. faba* (Aurora, ILB938 and Melodie) grouped with section *Hyperchusa* (*V. melanops*) species, although with very low bootstrap support. Additionally, there was a

monophyletic group composed of two subclades, the first subclade from section *pseudolathyrus* species *V. bithynica* and section *Ervilia* species *V. ervilia* with 33 % bootstrap whereas the second subclade was from members of section *Vicia*, namely *V. grandiflora* and *V. sativa* supported by 79 % bootstrap. This monophyly formed a paraphyly with all the above clades with 99% bootstrap support. Furthermore, section *Vicilla* members *V. amurensis* and *V. unijuga* had a separate clade that was supported by 50 % bootstrap.

Part of the big group section *Peregrinae* members *V. peregrina* and *V. michauxii* was monophyletic with species *V. sepium* from section *Sepium* with 38 % bootstrap support and their clade had 67 % bootstrap support, whereas *V. lathyroides* from section *Vicia* formed a paraphyletic grouping with both with the support of 27% bootstrap. Furthermore, from section *Cracca* including V. *ochroleuca* with *V. cracca* and then with *V. villosa*, formed a monophyletic clade with 94 % and 80 % bootstrap support, respectively. The last clade was composed of a mix of subgenus *Cracca* and subgenus *Vicia*. The monophyly from section *Narbonensis* such as *V. galilaea, V. narbonensis* and *V. serratifolia was* supported with 87 % bootstrap and they grouped with *V. hirsuta* from section *Cracca* with 26 % bootstrap support (Figure 5).

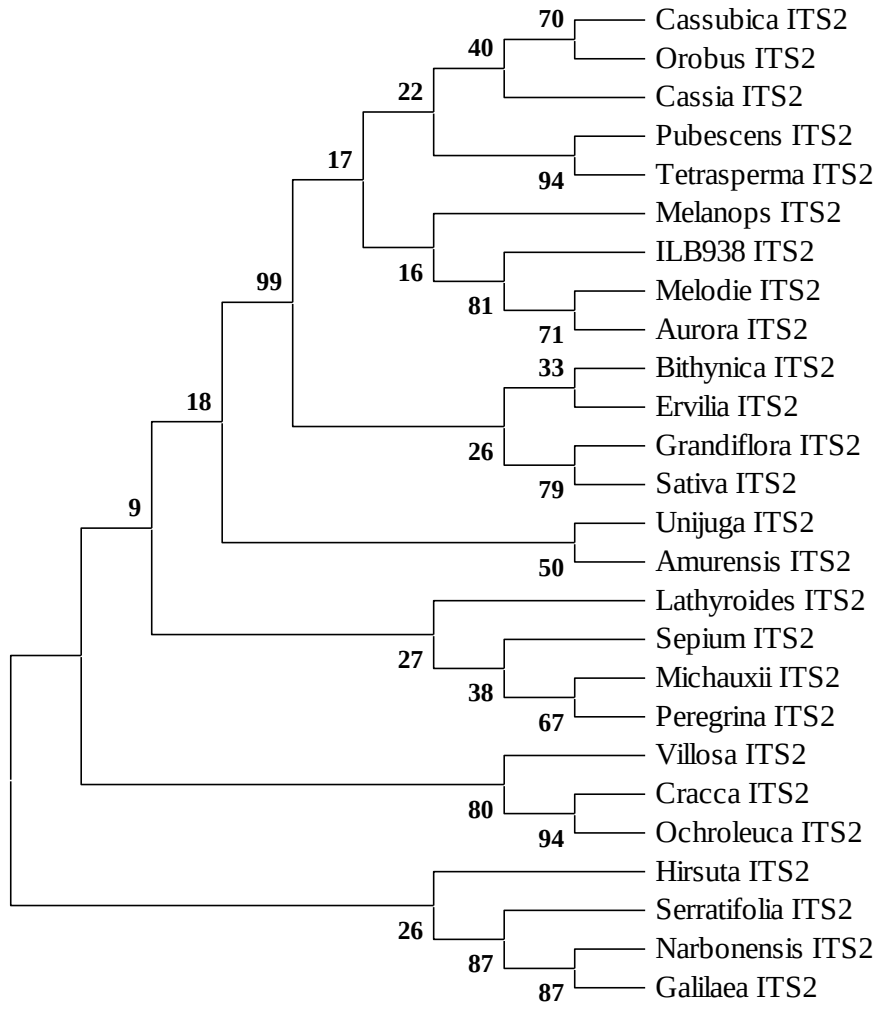

Figure 5. Phylogenies obtained from ML analysis of ITS2 sequence. Genus *Vicia* identified by their species name (Table1). The bootstrap consensus tree inferred from 1000 replicates. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. Phylogenetic analysis based ITS2 have two more accession of faba bean (ILB938 and Aurora).

5.2 Cox1 Phylogenetic Analyses

Model selection for 733 bp alignment of Cox1 sample resulted maximum likelihood fits substitution models. The lowest BIC scores were 2687.694. This analysis selected T92 (Tamura 3-parameter) for which maximum likelihood value (—ln L) = 1115.448, and nucleotide frequencies A and T=0.285, C and G=0.215. Rates of base substitutions were A—T =0.047, A—C, T—G and C—G =0.035, A—G and T—C = 0.144. Based on the selection of complete deletion, all positions containing gaps and missing data were eliminated. There was a total of 693 positions used for evolutionary analyses. Even though there was enough data to draw the phylogenetic tree, there was low divergence in the coding region, so it failed to distinguish between species. The major reason was the low rate of sequence change and conserved region. Therefore, there was no clear grouping that would help to draw a conclusion

about genus *Vicia* evolutionary family relationship. Almost all phylogenetic relationship resulted from Cox1 have very low (>50%) bootstrap support (Figure 6).

5.3 ITS2 and Cox1 Combined phylogenetic analyses

The concatenated sequence of ITS2 and Cox1 dataset had an average of 1247 bp in each of the 26 samples. According to the nucleotide substitution model, K2+G (Kimura 2-parameter: Gamma distribution) was the best fit with 3447.086 BIC score. The result included the maximum likelihood value (—ln L) is 1491.157 and the four nucleotide frequencies were 0.250. Rates of base substitutions of A—T and A—C =0.041, A—G and T—C = 0.167, T—G and C—G =0.041. The Gamma distribution parameter (+G) was 2.06 with 5 categories. All positions containing gaps and missing data were eliminated. There was a total of 349 positions in the final dataset that was used to analyze evolution. The phylogenetic tree from combined sequences of ITS2 and Cox1 had similar topologies with ITS2 phylogenetic tree except for some minor differences in their bootstrap support (Figure 7).

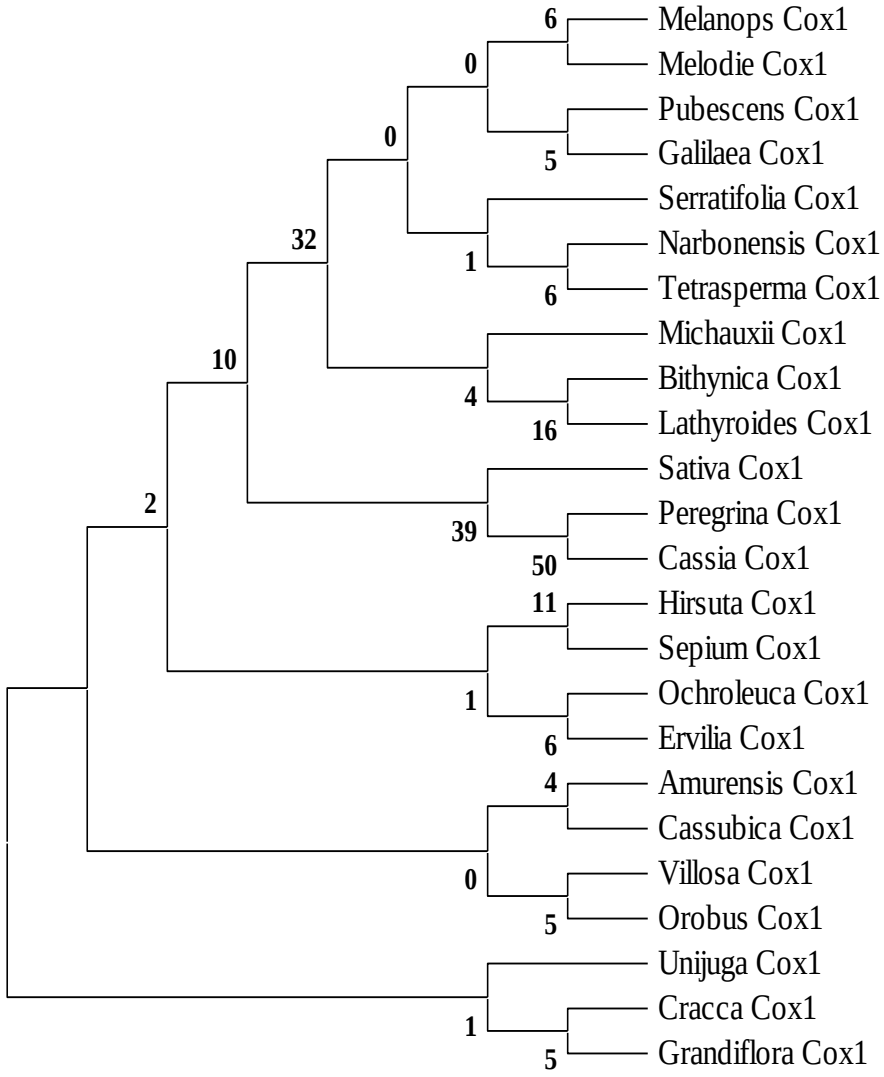

Figure 6. Genus *Vicia* phylogenetic tree based on Cox1 sequence. The bootstrap consensus tree inferred from 1000 replicates. Branches corresponding to partitions reproduced in less than 25% bootstrap replicates are collapsed. The species name and description of genus *Vicia* is as it is referred in Table1.

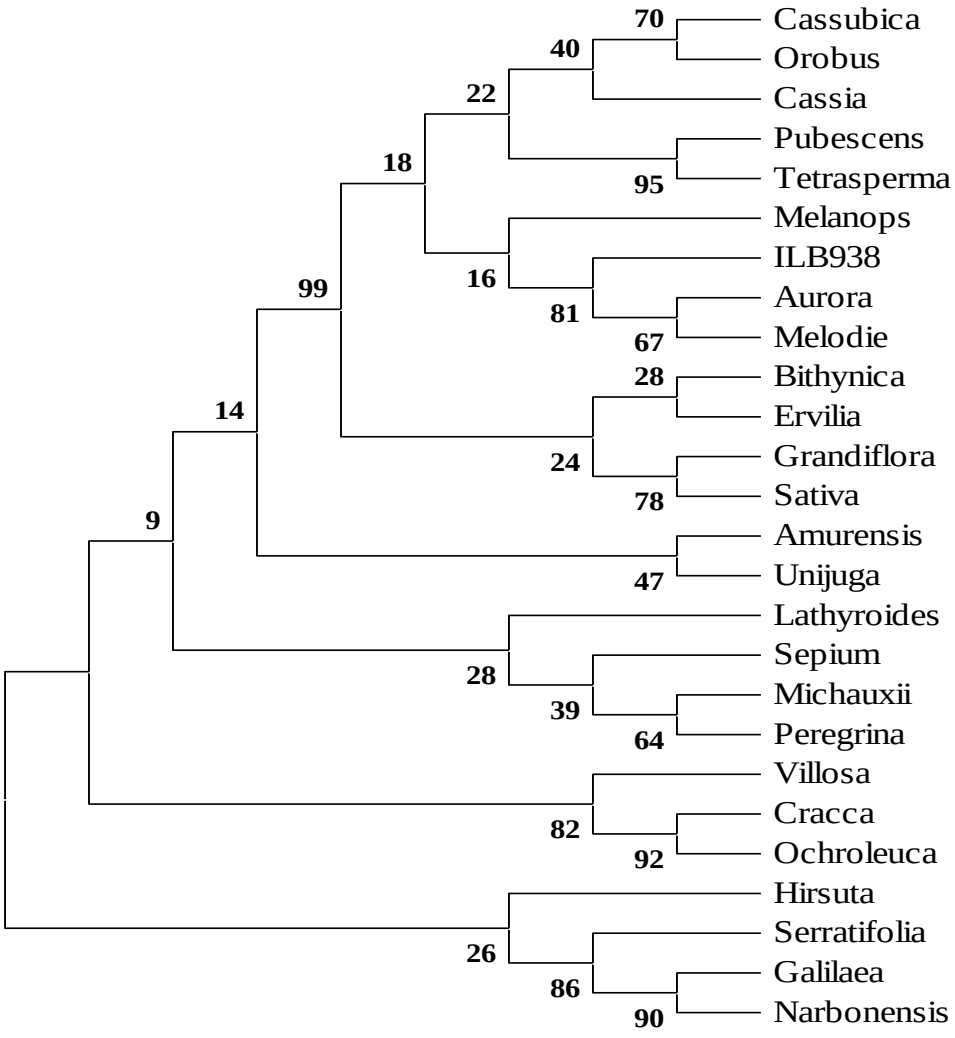

Figure 7. ITS2 and Cox1 combined phylogenetic analyses in Genus *Vicia*. The bootstrap consensus tree is inferred from 1000 replicates. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. They are described based on their species name which is listed in Table1.

6 Discussion

The aim of this study was to provide a molecular phylogeny of 24 species of genus *Vicia* from different locations in respect to how they group or how the clade is genetically related. For a better understanding of the phylogenetic relationship of any given plant family, the use of nuclear genes plays a pivotal role (Álvarez & Wendel, 2003). The success depends on their DNA sequence divergence among the target taxonomic group, and their ability to evolve and mutate in a certain time of a given evolution (Mugrabi de Kuppler, 2013).

The present results show that *V. faba* formed a monophyletic group with section *Hypechusa* (*V. Melanops*) and then to a group including sections *Ervum*, *Cassubicae*, and *Panduratae* but far from section *Narbonensis*. The clade of *V. faba* with *Hypechusa* (*V. Melanops*) is in line with Shiran et al., (2014) (Figure 7). However, the result contrasted with the work of Jaaska (1997) that stated a member of section *Narbonensis* forms close monophyly in a clade with species of *Hyperchusa* including *V. melanops*. On the other hand, In spite of morphological evidence, section *Faba* forms a separate grouping from a member of section *Narbonensis* with significant distance (Jaaska, 1997; Shiran et al., 2014). The cladistic and phenetic analysis of subgenus *Vicia* by Leht & Jaaska (2002) supported the inclusion of section *Faba* with section *Narbonensis* as did the morphological phylogenetic studied of Leht (2009).

Monophyletic grouping between *V. sativa* and *V. grandiflora* was supported with 78 % bootstrap, in line with Shiran et al., (2014). However, Jaaska (1997) as well as Leht & Jaaska (2002) grouped *V. grandiflora* in section *Sepium* together with *V. sepium* and *V. oroboides* rather than in the section *Vicia* with *V. sativa*. This group also forms a monophyletic group with section *Ervilia* (*V. ervilia*) and *Pseudolathyrus* (*V. bithynica*). This result contrasted with Leht & Jaaska (2002) as well as with Leht (2009) which states that *V. sativa* forms close groping with sections *Peregrinae*, *Hyperchusa*, *Pseudolathyrus* and *Narbonensis*. Except section *Pseudolathyrus*, it did not closely cluster with the rest of the listed sections (Figure 7).

Section *Ervilia* from subgenus *Cracca* and section *Pseudolathyrus* from subgenus *Vicia* formed one clade by their representative species *V. ervilia* and *V. bithynica*, respectively. Their monophyly relation had low bootstrap support (28%) (Figure 7). The separation of *V. ervilia* from other closely related sections such as *Ervoid* species is discussed in both Jaaska (2005) and Leht (2005) and supported by Endo et al. (2008). Neither of them suggested the probability of a monophyletic clade with another member of subgenus *Vicia*. *V. bithynica* was considered paraphyletic to section *Faba* and far from section *Narbonensis* as stated in Jaaska (1997). However, it is very distant from *V. lathyroides*.

Section *Cassubicae* including *V. cassubical* and *V. orobus* was in a clade with *V. cassia* from section *Panduralae*. Moreover, they had monophyletic relation to members of section *Ervum* such as *V. tetrasperma* and *V. pubescens* (Figure 7). The monophyly grouping of sections *Cassubicae* is the same result as Leht (2005) as well as the monophyletic of section *Cassubicae* supported by Leht (2009) if *V. amoena* is included in this section. Furthermore, Leht (2005) argued that section *Panduratae* should be part of section *Cracca*. However, in this result *V.*

cassia was very far from section *Cracca* (Figure 7). In addition, *Panduratae* as an independent section was supported by Jaaska (2005) and Leht (2009). Section *Ervum* member *V. tetrasperma* and *V. pubescen* were monophyletic with 92% bootstrap support and the results agreed with the work of Jaaska (2005) and Leht (2005), which stated section *Ervum* is far from the member of *Ervoid* species, so, it is not possible to form a separate unified subgenus. Regardless, Leht (2009) supported the separation of the *Ervoid* group to establish the third independent subgenus.

Section *Vicilla* related species *V. amurensis* and *V. unijuga* had a monophyly clade with average bootstrap support (Figure 7). They were polyphyletic to all the exceeding and basal clades. Hence, the inclusion of *V. amurensis* in section *Vicilla* is supported (Leht, 2005; Leht, 2009). However, the study of the phylogenetic significance of stylar features in genus *Vicia*, mentioned that both have a different stylar type (Choi et al., 2006).

The combination of *V. lathyroides, V. sepium,* and species of section *Peregrinae* such as *V. peregrina* and *V. michauxii* produced one clade (Figure 7). Jaaska (1997) stated that *V. lathyroides* formed subclades with sections *Sativa*, *Sepium*, and *Pseudolathyrus* by considering *Lathyroides* an independent section, though *V. lathyroides* formed a clade with *Sepium*. Still, *V. lathyroides* was distant from the rest of the other sections as well as from *V. faba*. Excluding that, they formed a monophyletic grouping with section *Peregrinae* (Leht & Jaaska, 2002; Leht, 2009). The separation of *V. lathyroides* from species of section *Vicia* is supported by Shiran et al., (2014). *V.sepium* and *V. grandiflora* were in a different clade of the phylogenetic tree with polyphyletic relation. According to Jaaska (1997), Leht & Jaaska (2002) and Leht (2009), V. *grandiflora* is supposed to be grouped with section *Sepium*. The monophyletic relation of section *Peregrinae* is well supported in the present results (Figure 7) as well as in Leht & Jaaska (2002); Leht, (2009). However, in the present phylogenetic tree, they did not form a linkage with sections *Hyperchusa,* instead they were very distant at low bootstrap support.

Section *Cracca* is one of the biggest sections from subgenus *Cracca* (Leht, 2005). In this study, *V. cracca, V. ochroleuca, and V. villosa* were used and they have monophyly grouping to each other (Figure 7*)*. The clade between *V. cracca* and *V. ochroleuca*, and theirs with *V. Villosa* is supported with 92% and 82% bootstrap, respectively. Furthermore, Leht (2005) placed the *V. cracca* under subsection apart from *V. villosa* whereas Jaaska, (2005) subgrouped *V. cracca* and *V. ochroleuca* together and another subgroup for *V. Villosa*. Additionally, Choi et al.

(2006) and Shiran et al. (2014) in their studies respectively mentioned that they shared common features and formed one group.

Species related to section *Narbonensis* such as *V. galilaea, V. narbonensis* and *V.serratifolia* had monophyly with *V. hirsuta.* Shiran et al. (2014) explained the grouping between sister taxa, the first grouping including *V. galilaea, V. serratifolia* and other species, while *V. narbonensis* formed a group with others. This contrasted with the present result, where the clade between *V. galilaea and V. narbonensis* had 90 % bootstrap support whereas monophyly relationship with *V. serratifolia* was supported by 86 % bootstrap. The clade between section *Narbonensis* and *V. hirsuta* was very weakly supported. Nevertheless, the separation of *V. hirsuta* from section *Cracca was* significantly supported (Leht, 2005; Choi et al., 2006; Endo et al., 2008; Leht, 2009; Shiran et al., 2014). Moreover, Jaaska, (2005) placed *V. hirsuta* in sections *Lenticula* as sister subclades with sections *Ervilia and Ervoides,* but this result has been neither supported nor refuted by any other researcher.

7 Conclusion

A molecular phylogenetic tree derived from combination of ITS2 and Cox1 proved to be useful to understand molecular evolution of species among genus *Vicia*. The present result was consistent with recently published work regarding the relation between *Vicia faba* and section *Narbonensis*, as it is stated that they are distantly related so that the placement of *V. faba* in different section is acceptable. Furthermore, based on maximum likelihood phylogeny analysis, it is possible to concluded *V. hirsuta* does not belong to section *Cracca*. Nevertheless, in which section *V. hirsuta* should be included must be examined in the future. In addition, *V. tetrasperma and V. pubescens* from section *Ervum* established monophyletic grouping with other member of subgenera *Vicia* and *Cracca* so that this work supported for the section to stay as part of subgenus *Cracca*. However, in order to give conclusive information on the phylogeny of genus *Vicia* the studies need to be carried in larger scale and also with different housekeeping gene since Cox1 was not effective as ITS2 to distinguish closely related species.

8 Acknowledgements

As my master's thesis project is a success, I would like to express my deepest appreciation to all those who provided me the opportunity and guidance. First of all, I thank the almighty God who assisted me in general. I also give a special gratitude to the Erasmus Mundus scholarship for financing me for the entire period of my study. I gratefully say thank you very much to my supervisor Adjunct Prof. Fred Stoddard who actively supported me in organizing and preparing this research project. I would also thanks to Laura Vottonen for her continuous support as well as Prof. Alan Schulman, Anne-Mari Narvanto and Wei Chang for their contribution. I would also like to acknowledge the important role of the staff of emPLANT program, especially Alicia Ayerdi-Gotor, Prof. Teemu Teeri, Aude Martin, Mélanie Bedot and Minna Haapalainen who got time to listen, guide, and keep me on the right path. Furthermore, distinguished appreciation also goes to my family and friends.

9 References

Abozeid, A., Liu, Y., Liu, J., & Tang, Z. (2018). Taxonomic implication of embryo micromorphology in the genus Vicia L. (Fabaceae). Plant Systematics and Evolution, 304(1), 33–42. https://doi.org/10.1007/s00606-017-1444-5

Alcoba-Flórez, J., Rivero-Pérez, B., & Méndez-Álvarez, S. (2007). Yeast molecular identification and typing. Communicating Current Research and Educational Topics and Trends in Applied Microbiology. Universidad de La Laguna, Tenerife, Spain. https://www.researchgate.net/publication/237227526_Yeast_molecular_identification_and_typing

Álvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution, 29(3), 417–434. https://doi.org/10.1016/S1055-7903(03)00208-2

Barkman, T. J., Chenery, G., McNeal, J. R., Lyons-Weiler, J., Ellisens, W. J., Moore, G., Wolfe, A. D., & dePamphilis, C. W. (2000). Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. Proceedings of the National Academy of Sciences of the USA, 97(24), 13166–13171. https://doi.org/10.1073/pnas.220427497

Berger, J. D., Robertson, L. D., & Cocks, P. S. (2003). Agricultural potential of Mediterranean grain and forage legumes: 2) Anti-nutritional factor concentrations in the genus Vicia. Genetic Resources and Crop Evolution, 50(2), 201–212. https://doi.org/10.1023/A:1022954232533

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Computational Biology, 15(4), e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Bowe, L. M., Coat, G., & dePamphilis, C. W. (2000). Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. Proceedings of the National Academy of Sciences of the USA, 97(8), 4092–4097. https://doi.org/10.1073/pnas.97.8.4092

Caputo, P., Frediani, M., Gelati, M. T., Venora, G., Cremonini, R., & Castiglione, M. R. (2013). Karyological and molecular characterisation of subgenus Vicia (Fabaceae). Plant Biosystems - 147(4), 1242–1252. https://doi.org/10.1080/11263504.2013.861532

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2016). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. Molecular Ecology Resources, 16(1), 138–149. https://doi.org/10.1111/1755-0998.12438

Choi, B.-H., Seok, D.-I., Endo, Y., & Ohashi, H. (2006). Phylogenetic significance of stylar features in genus Vicia (Leguminosae): An analysis with molecular phylogeny. Journal of Plant Research, 119(5), 513–523. https://doi.org/10.1007/s10265-006-0015-6

Cowan, R. S., Chase, M. W., Kress, W. J., & Savolainen, V. (2006). 300,000 species to identify: Problems, progress, and prospects in DNA barcoding of land plants. TAXON, 55(3), 611–616. https://doi.org/10.2307/25065638

Endo, Y., Choi, B.-H., Ohashi, H., & Delgado-Salinas, A. (2008). Phylogenetic Relationships of New World Vicia (Leguminosae) Inferred from nrDNA Internal Transcribed Spacer Sequences and Floral Characters. Systematic Botany, 33(2), 356–363. https://doi.org/10.1600/036364408784571536

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M., & Barrett, S. C. H. (2008a). Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. PLoS ONE, 3(7), e2802. https://doi.org/10.1371/journal.pone.0002802

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39(4), 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Frediani, M., Maggini, F., Gelati, M. T., & Cremonini, R. (2004). Repetitive DNA sequences as probes for phylogenetic analysis in Vicia genus. Caryologia, 57(4), 379–386. https://doi.org/10.1080/00087114.2004.10589420

Gold and Green (2020) https://goldandgreenfoods.com/sustainability/health-and-wellbeing/, viewed 7.5.2021

Hanelt, P., & Mettin, D. (1989). Biosystematics of the Genus Vicia L. (Leguminosae). Annual Review of Ecology and Systematics, 20, 199-223. https://www.jstor.org/stable/2097090?seq=1

Jaaska, V. (1997). Isoenzyme diversity and phylogenetic affinities in Vicia subgenus Vicia (Fabaceae). Genetic Resources and Crop Evolution, 44(6), 557–574. https://doi.org/10.1023/A:1008630003045

Jaaska, V. (2005). Isozyme Variation and Phylogenetic Relationships in Vicia subgenus Cracca (Fabaceae). Annals of Botany, 96(6), 1085–1096. https://doi.org/10.1093/aob/mci260

Janssens, S., Couvreur, T. L. P., Mertens, A., Dauby, G., Dagallier, L.-P., Vanden Abeele, S., Vandelook, F., Mascarello, M., Beeckman, H., Sosef, M., Droissart, V., van der Bank, M., Maurin, O., Hawthorne, W., Marshall, C., Réjou-Méchain, M., Beina, D., Baya, F., Merckx, V., Hardy, O. (2020). A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses. Biodiversity Data Journal, 8, e39677. https://doi.org/10.3897/BDJ.8.e39677

Knoop, V. (2004). The mitochondrial DNA of land plants: Peculiarities in phylogenetic perspective. Current Genetics, 46(3), 123–139. https://doi.org/10.1007/s00294-004-0522-8

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005a). Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences of the USA, 102(23), 8369–8374. https://doi.org/10.1073/pnas.0503123102

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution, 35(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lahuta, L. B., Ciak, M., Rybiński, W., Bocianowski, J., & Börner, A. (2018). Diversity of the composition and content of soluble carbohydrates in seeds of the genus Vicia (Leguminosae). Genetic Resources and Crop Evolution, 65(2), 541–554. https://doi.org/10.1007/s10722-017-0552-y

Leht, M. (2005). Cladistic and phenetic analyses of relationships in Vicia subgenus Cracca (Fabaceae) based on morphological data. Taxon, 54(4), 1023–1032. https://doi.org/10.2307/25065486

Leht, M. (2009). Phylogenetics of Vicia (Fabaceae) based on morphological data. Feddes Repertorium, 120(7–8), 379–393. https://doi.org/10.1002/fedr.200911117

Leht, M., & Jaaska, V. (2002). Cladistic and phenetic analysis of relationships in Vicia subgenus Vicia (Fabaceae) by morphology and isozymes. Plant Systematics and Evolution, 232(3–4), 237–260. https://doi.org/10.1007/s006060200045

Micales, J. A., & Bonde, M. R. (2017). Isozymes: Methods and Applications *. In R. P. Singh & U. S. Singh (Eds.), Molecular Methods in Plant Pathology (1st ed., pp. 115–130). CRC Press. https://doi.org/10.1201/9780203746523-9

Miranda, V. F. O. de, Martins, V. G., Furlan, A., & Bacci Jr., M. (2010). Plant or fungal sequences? An alternative optimized PCR protocol to avoid ITS (nrDNA) misamplification. Brazilian Archives of Biology and Technology, 53(1), 141–152. https://doi.org/10.1590/S1516-89132010000100018

Mugrabi de Kuppler, A. L. (2013). Phylogenetics, flow-cytometry, and pollen storage in Erica L. (Ericaceae). Implications for plant breeding and interspecific crosses. (Doctoral dissertation, Bonn university,Germany). https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/5565

Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York. https://doi.org/10.1016/S1055-7903(02)00245-2.

Poczai, P., & Hyvönen, J. (2010). Nuclear ribosomal spacer regions in plant phylogenetics: Problems and prospects. Molecular Biology Reports, 37(4), 1897–1912. https://doi.org/10.1007/s11033-009-9630-3

Potokina, E., Tomooka, N., Vaughan, D. A., Alexandrova, T., & Xu, R.-Q. (1999). Phylogeny of Vicia Subgenus Vicia (Fabaceae) Based on Analysis of RAPDs and RFLP of PCR-Amplified Chloroplast Genes. Genetic Resources and Crop Evolution, 46(2), 149–161. https://doi.org/10.1023/A:1008640322198

Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. Journal of Plant Ecology, 9(2), 233–239. https://doi.org/10.1093/jpe/rtv047

Raina, S. N., & Ogihara, Y. (1995). Ribosomal DNA repeat unit polymorphism in 49 Vicia species. Theoretical and Applied Genetics, 90(3–4), 477–486. https://doi.org/10.1007/bf00221993

Raveendar, S., Lee, J.-R., Park, J.-W., Lee, G.-A., Jeon, Y.-A., Lee, Y. J., Cho, G.-T., Ma, K.-H., Lee, S.-Y., & Chung, J.-W. (2015). Potential use of ITS2 and matK as a Two-Locus DNA Barcode for Identification of Vicia Species. Plant Breeding and Biotechnology, 3(1), 58–66. https://doi.org/10.9787/PBB.2015.3.1.058

Roze I., Rūrāne I. 2013. Revision of genus Vicia L. in the flora of Latvia. Acta Biol. Univ. Daugavp., 13 (1): 85 – 93.

Shiran, B., Kiani, S., Sehgal, D., Hafizi, A., ul-Hassan, T., Chaudhary, M., & Raina, S. N. (2014). Internal transcribed spacer sequences of nuclear ribosomal DNA resolving complex taxonomic history in the genus Vicia L. Genetic Resources and Crop Evolution, 61(5), 909–925. https://doi.org/10.1007/s10722-014-0086-5

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22(22), 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Tippery, N. P., & Les, D. H. (2008). Phylogenetic analysis of the internal transcribed spacer (ITS) region in Menyanthaceae using predicted secondary structure. Molecular Phylogenetics and Evolution, 49(2),526–537. https://doi.org/10.1016/j.ympev.2008.07.019

van de Wouw, M., Enneking, D., Robertson, L. D., & Maxted, N. (2001). Vetches (Vicia L.). In N. Maxted & S. J. Bennett (Eds.), Plant Genetic Resources of Legumes in the Mediterranean, 134–158. Springer Netherlands. https://doi.org/10.1007/978-94-015-9823-1_8

Verso Food (2021) https://www.beanit.fi/en/ viewed 7.5.2021

Wojciechowski, M. F., Lavin, M., & Sanderson, M. J. (2004). A phylogeny of legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. American Journal of Botany, 91(11), 1846–1862. https://doi.org/10.3732/ajb.91.11.1846

Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., & Wang, G. T. (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources, 20(1), 348–355. https://doi.org/10.1111/1755-0998.13096